AD-A034 064 MASSACHUSETTS INST OF TECH CAMBRIDGE DEPT OF ELECTRI--ETC F/G 12/2
THE USE OF GROUP-THEORETIC CONCEPTS IN SOLVING PROBLEMS IN ESTI--ETC(U)
1976 A S WILLSKY AF-AFOSR-2273-72

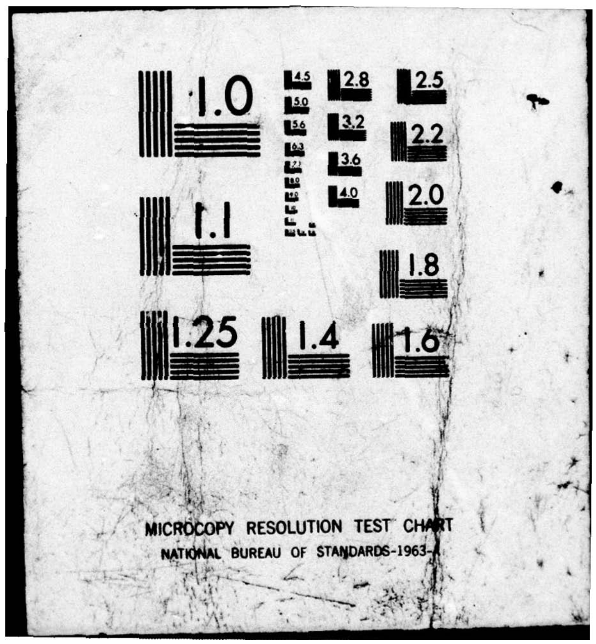
UNCLASSIFIED AFOSR-TR-76-1309 NL

| OF |

AD
A034064

END

DATE
FILMED
2-77


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
4.5
5.0
5.6
6.3
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

AFOSR - TR - 76 - 1309

THE USE OF GROUP-THEORETIC CONCEPTS IN SOLVING PROBLEMS IN ESTIMATION AND CONTROL

Alan S. Willsky*

Electronic Systems Laboratory
Department of Electrical Engineering and Computer Science
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

Abstract

In this paper we describe some results obtained over the past few years in the areas of estimation and control. The point of this research is to find subclasses of systems that possess sufficient structure to allow us to determine optimal solutions or computationally feasible algorithms. In particular, we utilize several results from group theory (both for Lie and finite groups) to obtain our results.

I. Introduction

Although in principle many control and estimation problems have been solved in extremely general settings, the solutions are often unacceptable because

(a) The "solution" does not actually provide a design (as in the so-called Kushner-Stratonovich solution of the nonlinear filtering problem).

(b) The resulting design is much too complex to be evaluated for a specific system (as in solving extremely high-order) Riccati equations).

or (c) The design is computationally too complex for on-line implementation (as with many nonlinear filtering algorithms or the control of large interconnected systems).

Often these difficulties arise because of the extreme generaltiy of the problem being solved. The approach of the research described in this paper has been to restrict attention to subclasses of systems and problems in order to exploit the algebraic structure that they have in common. Our motivation has been that there must be a middle ground between "linear" and "nonlinear" or between "small-scale" and "large scale" -- i.e. that while in both cases the former categories may be too restrictive for many applications, the latter are far too large to provide many useful results. In this paper we describe several results--all using group-theoretic concepts--for classes of systems between these structural extremes. It is our feeling that results of this type not only provide useful solutions to certian problems but also shed considerable light on the implications of structure. That is, we can begin to answer questions such as which results really require linearity and which can we extend to systems with somewhat less algebraic structure.

II. Finite-Dimensional Optimal Nonlinear Filters

The results we review in this section were obtained in collaboration with James Lo and Steven Marcus. For thorough treatments of these and related results and for further references, we refer the reader to [1-8]. Consider a dynamical system described by the following equations:

$$\dot{x}(t) = \left[A_0 + \sum_{i=1}^{N} A_i \xi_i(t)\right] x(t) \qquad (2.1)$$

$$x(0) = x_0 \quad \text{(given)}$$

$$d\xi(t) = F\xi(t)dt + Gdw(t) \qquad (2.2)$$

where $\xi \in R^n$, $x \in R^k$ or R^{kxk}, and w is a standard Brownian motion process. Such a system is often called a <u>bilinear system</u>, and systems of this type arise in a wide variety of applications (see [1-3,5-7]). One very important application arises when $x(t)$ is a 3x3 <u>direction cosine matrix</u>, representing the orientation of a rigid body with respect to inertial space (see [6,7] for details). Here the A_i are all skew-symmetric and the ξ_i represent components of the angular velocity of the rigid body (in this case, of course, an equation of the form (2.2) holds only for spherically-symmetric bodies; otherwise the equations contain nonlinearities, reflecting differences in the body's moments of inertia).

The estimation problem we wish to consider is as follows: we observe

$$dz(t) = H\xi(t)dt + R^{1/2}dv(t) \qquad (2.3)$$

where $R > 0$ and v is a standard Brownian motion process independent of w. Given $z(s)$, $0 \leq s \leq t$, we want to devise a recursive procedure for esti-

---

* Associate Professor of Electrical Engineering. This work was supported by AFOSR under Grant 72-2273.

AFOSR - TR - 76 - 1309

mating $x(t)$ -- i.e. we wish to compute

$$\hat{x}(t|t) = E[x(t)|z(s),\ 0\leq s\leq t] \qquad (2.4)$$

Problems of this type arise in rigid body attitude estimation problems, when we have a strapdown system that provides measurements of angular velocity (see [6]) or when we have an inertial platform (see [6,7]).

In order to understand this problem let us consider the simplest case -- N=n=k=1. In this case

$$x(t) = \left[e^{A_0 t}\, e^{A_1 \int_0^t \xi(s)ds}\right] x_0 \qquad (2.5)$$

The solution then is clear: let

$$\rho(t) = \int_0^t \xi(s)ds \qquad (2.6)$$

and design a Kalman filter to compute the optimal estimate $(\hat{\xi}(t|t), \hat{\rho}(t|t))$ of the augmented state $(\xi(t),\rho(t))$. Then

$$\hat{x}(t|t) = e^{A_0 t} E[e^{A_1\rho(t)} | z(s),\ 0\leq s\leq t]x_0 \qquad (2.7)$$

and, since ρ is Gaussian when conditioned on z, we can compute $\hat{x}(t|t)$ as a function of $\hat{\rho}(t|t)$ and its covariance

$$\hat{x}(t|t) = e^{A_0 t + A_1\hat{\rho}(t|t) + \frac{1}{2}A_1^2\sigma(t|t)} \qquad (2.8)$$

$$\sigma(t|t) = E[(\rho(t)-\hat{\rho}(t|t))^2] \qquad (2.9)$$

where σ can be precomputed.

The key question is how this simple result extends in the multivariable case. Note that if all of the A_i commute, we can write

$$\hat{x}(t) = \exp\left\{A_0 t + \sum_{i=1}^{N} A_i \int_0^t \xi_i(S)ds\right\} x_0 \qquad (2.10)$$

and in this case one can readily extend the scalar result (see [1]). When the A_i don't commute the situation is far more complex, and, in fact, in order to obtain results, one must determine the manner in which the A_i don't commute. This leads to the introduction of some notions from the theory of Lie groups and Lie algebras.

Definitions (1) Let L be a subspace of kxk matrices (over C). L is a Lie algebra if

$$[A,B] \triangleq AB-BA \in L \qquad \forall A,B\in L \qquad (2.11)$$

(2) Let S be a set of kxk matrices. Then there exists a smallest Lie algebra $\{S\}_A$ that contains S. This is called the algebra generated by S.

(3) Let L be a Lie algebra. The Lie group G associated with L is the smallest group (under matrix multiplication) containing $\exp(L)$ $\forall L\in L$.

(4) Let L be a Lie algebra. A subalgebra S is an ideal if

$$[L,S] \triangleq \{[L,S]|L\in L,\ S\in S\} \subseteq S \qquad (2.12)$$

If S is a set in L, we let $\{S\}_{I,L}$ denote the ideal in L generated by S.

(5) A Lie algebra L is nilpotent if the lower central series of ideals

$$L^0 = L$$
$$L^{n+1} = [L,L^n] \qquad (2.13)$$

terminates in $\{0\}$ for some n. Note that L is abelian if and only if $L^1 = \{0\}$.

We refer the reader to [1,5,6] and in particular to [9,10] for physical and mathematical motivation for the introduction of Lie-theoretic concepts into the problem of analysis of bilinear systems. We only note the following: we associate two Lie algebras with the system (2.1)

$$L = \{A_0,A_1,\ldots,A_N\}_A$$
$$L_0 = \{A_1,\ldots,A_N\}_{I,L} \qquad (2.14)$$

Then if G_0 is the Lie group associated with L_0,

$$x(t) \in e^{At}\, G_0\, x_0 \qquad (2.15)$$

For more detailed controllability results, we refer the reader to [9,10].

We have already stated that if L is abelian, we can obtain a finite-dimensional optimal estimator for $x(t)$, and in fact that estimator consists of a Kalman filter followed by an exponentiation. More generally, we obtain the result [6,8].

Theorem 2.1: Suppose L_0 is nilpotent. Then $\hat{x}(t|t)$ can be computed via a finite set of (in general nonlinear) stochastic differential equations (driven by the innovations $d\nu(t) = dz(t) - H\hat{\xi}(t|t)$).

The basis for this result is as follows: if L_0 is nilpotent, then one can readily show [6,8] that every element of $x(t)$ is a finite sum of terms of the form

$$\phi(t) = e^{\int_0^t c'\xi(s)ds}\, \eta(t) \qquad (2.16)$$

where $\eta(t)$ is a finite Volterra series

$$\eta(t) = w_0(t) + \sum_{j=1}^{N} \int_0^t \cdots \int_0^t \sum_{k_1,\ldots,k_j=1}^{N} w_j^{(k_1,\ldots,k_j)}(t,\sigma_1,\ldots,\sigma_j)\, \xi_{k_1}(\sigma_1)\cdots\xi_{k_j}(\sigma_j)\,d\sigma_1\cdots d\sigma_j$$

where the kernels are separable - i.e. the w's are of the form

$$w(t,\sigma_1,\ldots,\sigma_j) = \sum_{i=1}^{m} \gamma_0^i(t)\gamma_1^i(\sigma_1)\cdots\gamma_j^i(\sigma_j) \quad (2.18)$$

Theorem 2.1 then follows from the more general.

Theorem 2.2 ([6,8]): The estimate $\hat{\phi}(t|t)$ can be computed with a finite-dimensional set of stochastic differential equations if η is given by (2.17) with all separable kernels. This result includes far more system than those described by (2.1). We refer the reader to [6,8,11] for examples and more on systems with such Volterra series representations.

The proof of this result is quite lengthy and complicated, and we refer the reader to [6,8] for details. We indicate only a few of the critical ideas. The separability condition is essential to have finite-dimensional realizability. Consider one of the simplest cases

$$\phi(t) = \int_0^t w(t,s)\xi(s)ds \quad (2.19)$$

Using a conditional version of Fubini's theorem, we have

$$\hat{\phi}(t|t) = \int_0^t w(t,s)\hat{\xi}(s|t)ds \quad (2.20)$$

At first glance, it appears that we have to smooth the entire ξ trajectory in order to estimate ϕ. However, if w is separable, then we can realize (2.19) by a finite-dimensional linear systems which when combined with (2.2) leads to a finite-dimensional Kalman filter that can compute (2.20). The situation for higher order Volterra terms is more complicated but similar.

Now define the conditional smoothed covariance

$$P_{ij}(\sigma_1,\sigma_2,t) = E[(\xi_i(\sigma_1)-\hat{\xi}_i(\sigma_1|t))(\xi_j(\sigma_2)-\hat{\xi}_j(\sigma_2|t))] \quad (2.21)$$

A critical property of P is that it is separable. To see the importance of this consider

$$\phi(t) = \int_0^t \int_0^{\sigma_1} \int_0^{\sigma_2} \xi_1(\sigma_1)\xi_2(\sigma_2)\xi_3(\sigma_3)d\sigma_3 d\sigma_2 d\sigma_1 \quad (2.22)$$

In this expression, write

$$\xi_i(\sigma_i) = [\xi_i(\sigma_i)-\hat{\xi}_i(\sigma_i|t)] + \hat{\xi}_i(\sigma_i|t) \quad (2.23)$$

One of the terms that results from this is

$$\psi(t) = \int_0^t \int_0^{\sigma_1} \int_0^{\sigma_2} [\xi_1(\sigma_1)-\hat{\xi}_1(\sigma_1|t)][\xi_2(\sigma_2)-\hat{\xi}_2(\sigma_2|t)]\hat{\xi}_3(\sigma_3|t)d\sigma_3 d\sigma_2 d\sigma_1$$

Then it is easily shown that

$$\hat{\psi}(t|t) = \int_0^t \int_0^{\sigma_1} \int_s^{\sigma_2} P_{12}(\sigma_1,\sigma_2,t)\hat{\xi}_3(\sigma_3|t)d\sigma_3 d\sigma_2 d\sigma_1$$

$$- \int_0^t \int_0^{\sigma_1} P_{12}(\sigma_1,\sigma_2,t)\hat{\rho}_3(\sigma_2|t)d\sigma_2 d\sigma_1 \quad (2.24)$$

where

$$\rho_3(\sigma) = \int_0^{\sigma} \xi_3(s)ds$$

Again, it is easily seen that the separability of P_{12} allows us to augment ξ with a finite set of variables in order to determine a Kalman filter that computes $\hat{\psi}(t|t)$.

One of the other terms in (2.22), (2.23) is

$$\mu(t) = \int_0^t \int_0^{\sigma_1} \int_0^{\sigma_2} \hat{\xi}_1(\sigma_1|t)\hat{\xi}_2(\sigma_2|t)\hat{\xi}_3(\sigma_3|t)d\sigma_3 d\sigma_2 d\sigma_1$$

Taking the differential we obtain

$$d\mu(t) = \hat{\xi}_1(t|t)\lambda(t)dt + \int_0^t \int_0^{\sigma_1} \int_0^{\sigma_2} d[\hat{\xi}_1(\sigma_1|t)\hat{\xi}_2(\sigma_2|t)\hat{\xi}_3(\sigma_3|t)]d\sigma_3 d\sigma_2 d\sigma_1 \quad (2.26)$$

$$\lambda(t) = \int_0^t \int_0^{\sigma_1} \hat{\xi}_2(\sigma_1|t)\hat{\xi}_3(\sigma_2|t)d\sigma_2 d\sigma_1 \quad (2.27)$$

Using Ito's differential rule on the second term in (2.26), computing $d\lambda(t)$ in much the same manner, and iterating these arguments, we obtain a finite-dimensional set of nonlinear differential equations *(driven by the innovations) for the computation* of μ (see [6]). The general result follows from arguments along these lines and from related properties of conditional Gaussian processes.

Thus, we see that this result combines some relatively simple properties of linear filters with the algebraic structure of nilpotent bilinear systems in order to provide one of the largest classes of finite dimensional optimal nonlinear estimators that in some sense can be viewed as a "natural" extension of the Kalman filter [1,6].

3. Fast Estimation Algorithms for Certain Markov Chains

Let $x(k)$ be a finite state Markov process with state set $S = \{s_1,\ldots,s_N\}$ and with probability vector $p(k)$ (here $p(k)_i = \text{Prob}\{x(k)=s_i\}$). The probability transition equation for the process is

$$p(k+1) = Ap(k) \tag{3.1}$$

where A is an NxN stochastic matrix. Suppose we take a sequence $\{y(k)\}$ of noisy measurements of the process. These measurements are characterized by the conditional distribution (if y is discrete) or density (if y is continuous)

$$C(\mu,i) = P(y(k)=\mu|x(k)=s_i) \tag{3.2}$$

we also assume that, *conditioned on* $x(k)$, $y(k)$ is independent of all other $x(j)$.

The problem of computing the conditional distributions $p(k|k)$ and $p(k+1|k)$ for $x(k)$ and $x(k+1)$ conditioned on $y(1),\ldots,y(k)$ is conceptually trivial, and the solution is given by

Diffusion Update

$$p(k+1|k) = Ap(k|k) \tag{3.3}$$

Measurement Update

$$p(k|k)_i = \frac{C(y(k),i)p(k|k-1)_i}{\sum_{j=1}^{N} C(y(k)j)p(k|k-1)_j} \tag{3.4}$$

Although the derivation of (3.3), (3.4) is a trivial application of Bayes' rule, the implementation of these equations can be computationally nontrivial. For example, in general we must perform N^2 multiplications in carrying out the diffusion update. The question we wish to consider in this section is the investigation of special classes of finite-state Markov process for which the estimation procedure (3.3), (3.4) can be performed efficiently. The results we discuss in this section and other related topics are treated in much more detail in [12-15].

To motivate the approach, suppose $x(k) \varepsilon Z_N$ = the group of integers $\{0,1,\ldots,N-1\}$ with addition defined modulo N. Furthermore, suppose

$$x(k+1) = [x(k) + w(k)] \bmod N \tag{3.5}$$

where w is a sequence of identically distributed, independent, random variables with values in Z_N and with distribution ξ. We also assume

$$y(k) = [x(k) + v(k))] \bmod N \tag{3.6}$$

where v is a sequence of independent Z_N -variables, independent of w. Let $\eta(k)$ be the distribution for $-v(k)$ (here "minus" is interpreted modulo N). In this case, several straightforward calculations yield the filtering algorithm

$$p(k+1|k)_i = \sum_{j=0}^{N-1} \xi_j p(k|k)_{i-j} \tag{3.7}$$

$$p(k|k)_i = \frac{\eta(k)_{i-y(k)}p(k|k-1)_i}{\sum_{j=0}^{N-1} \eta(k)_{j-y(k)}p(k|k-1)_j} \tag{3.8}$$

where all indices are to be interpreted mod N.

Let us identify Z_N with $\{1,z,z^2,\ldots,z^{N-1}\}$ where z is an indeterminate and we multiply with the rule $z^N = 1=z^0$ (hence this is a group isomorphic to Z_N). Then any function $f: Z_N \to C$ is represented in the form

$$F = \sum_{i=0}^{N-1} f_i z^i \tag{3.9}$$

We equip the set of all such functions with the usual definitions of pointwise addition and scalar multiplication and with a convolution product

$$f*g = \left(\sum_{i=0}^{N-1} f_i z^i\right)*\left(\sum_{j=0}^{N-1} g_j z^j\right) = \sum_{i,j=0}^{N-1} f_i g_j z^{i+j} = \sum_{t=0}^{N-1} \gamma_t z^t \tag{3.10}$$

where

$$\gamma_t = \sum_{\ell=0}^{N-1} f_\ell g_{t-\ell} \tag{3.11}$$

This set, denote by $C[Z_N]$ is called the (complex) group algebra of Z_N , and, with this notation (3.7) becomes

$$p(k+1|k) = \xi * p(k|k) \tag{3.12}$$

If we also equip $C[Z_N]$ with the pointwise product

$$(fg)_i = f_i g_i \tag{3.13}$$

and the evaluation map

$$s(f) = \sum_{i=0}^{N-1} f_i \tag{3.14}$$

then (3.8) becomes

$$\lambda(k) = \eta * y(k) \tag{3.15}$$

$$\gamma(k|k) = \lambda(k) p(k|k-1) \tag{3.16}$$

$$p(k|k) = \frac{\gamma(k|k)}{s[\gamma(k|k)]} \tag{3.17}$$

(here we regard $y(k)$ as an element of $C[Z_N]$).

Examining these equations we see that the diffusion update (3.12) consists of a cyclic convolution, while the measurement update consists of a permutation (3.15), a pointwise multiplication (3.16), and a normalization (3.17). Note that if we wrote (3.12) in the vector form (3.3), A would be a circulant matrix.

The presence of a cyclic convolution in these equationssuggests the possible use of transforms. For $f \in C[Z_N]$, define

$$c_i(f) = \frac{1}{N} \sum_{m=0}^{N-1} f_m \gamma^{-im} \tag{3.18}$$

where

$$\gamma = e^{j2\pi/n}. \tag{3.19}$$

Note the inverse transform relation

$$f_m = \sum_{i=0}^{N-1} c_i(f) \gamma^{im} \tag{3.20}$$

Applying transforms to the algorithm (3.12), (3.15)-(3.17) yields

$$c_i(p(k+1|k)) = Nc_i(\xi) c_i(p(k|k)) \tag{3.21}$$

$$c_i(\lambda(k)) = \gamma^{-iy(k)} c_i(\eta) \tag{3.22}$$

$$c_i(\gamma(k|k)) = \sum_{j=0}^{N-1} c_j(\lambda(k)) c_{i-j} \tag{3.23}$$

$$c_i(p(k|k)) = \frac{c_i(\gamma(k|k))}{Nc_0(\gamma(k|k))} \tag{3.24}$$

Examining these equations we note two things:

(1) Comparing (3.12), (3.16) with (3.21), (3.23) we see a striking "duality". In the $C[Z_N]$ domain the diffusion update is a convolution and the central step of the measurement update is a pointwise product. On the other hand, in the transform domain the situation is exactly reversed.

(2) We can exploit this duality to reduce the overall computational complexity from $O(N^2)$ to $O(N \log N)$ with the aid of the fast Fourier transform (FFT).

Motivated by these observations, we consider a more general problem (for a far more general case, see [14-15]). To do this we need the following ([16])

Definitions: (1) Let X be a group. The (complex) group algebra $C[X]$ consists of all formal sums

$$f = \sum_{g \in X} f_g \cdot g \qquad f_g \in C \tag{3.25}$$

with pointwise addition and scalar multiplication, and the convolution product

$$f * d = \sum_{g,h \in X} f_g d_h \cdot gh = \sum_{t \in X} \gamma_t \cdot t \tag{3.26}$$

$$\gamma_t = \sum_{g \in X} f_g d_{g^{-1}t} \tag{3.27}$$

We also endow $C[X]$ with the pointwise product and the evaluation map

$$(fg)_h = f_h g_h \quad , \quad s(f) = \sum_{g \in X} f_g \tag{3.28}$$

(2) A representation T of X is a homomorphism of $T: X \to C^{k+k}$

(3) Such a representation is called irreducible if the only subspaces V of C^k such that

$$T(x) V \subset V \qquad \forall x \in X \tag{3.29}$$

are $\{0\}$ and C^k.

(4) Two representations T_1 and T_2 are equivalent if there exists an invertible matrix S such that

$$T_1(x) = ST_2(x)S^{-1} \qquad \forall x \in X \tag{3.30}$$

Let X be a given finite group with n elements. Then [16] one can find a complete set of inequivalent.irreducible representations $T^1, \ldots, T^s$ with $\dim T^i = z^i$, $T^1 \equiv 1$, and

$$n = \sum_{i=1}^{s} z_i^{\,2} \tag{3.31}$$

Furthermore, one has the transform pair

$$c^i(\phi) = \frac{z_i}{n} \sum_{g \in X} \phi_g [T^i(g^{-1})]' \tag{3.32}$$

$$\phi_g = \sum_{i=1}^{s} \sum_{j,k=1}^{z_i} c^i_{jk}(\phi) T^i_{jk}(g) \tag{3.33}$$

Consider a random process $x(k)$ and a sequence of observations $y(k)$ on X described by

$$x(k+1) = w(k)x(k) \tag{3.34}$$

$$y(k) = v(k)x(k) \tag{3.35}$$

where w and v are independent sequences of independent random variables on X. Let ξ be the distribution of w, and let η be the distribution of v^{-1}. In the group algebra domain we obtain the following filtering algorithm

$$\rho(k+1|k) = \xi * \rho(k|k) \tag{3.36}$$

$$\lambda(k) = \eta * y(k) \tag{3.37}$$

$$\gamma(k|k) = \lambda(k)\rho(k|k-1) \tag{3.38}$$

$$\rho(k|k) = \frac{\gamma(k|k)}{s[\gamma(k|k)]} \tag{3.39}$$

If we consider the transform domain version of the algorithm we find

$$C^i(\rho(k+1|k)) = C^i(\xi)C^i(\rho(k|k)) \tag{3.40}$$

$$C^i(\lambda(k))) = C^i(\eta)[T^i(y(k)^{-1})]' \tag{3.41}$$

$$C^i(\rho(k|k)) = \frac{C^i(\gamma(k\ k))}{nC^1(\gamma(k|k))} \tag{3.42}$$

and we can make the following remarks:

(1) The convolution (3.36) of the diffusion update is transformed into a pointwise multiplication (3.40) of the transform matrices.

(2) The transform version of the pointwise multiplication (3.38) is in general quite complex (see [15]). However, in certain cases -- e.g. when X is a metacyclic group [16] -- it can be interpreted as a convolution of the transforms, yielding the "duality" encountered earlier in the Z_N case.

(3) In order to take advantage of pointwise multiplication versus a convolution, we need to be able to perform the transforms (3.32), (3.33) quickly. A general procedure for this is not yet known, but in certain cases (such as the metacyclic case) we can obtain generalizations of the FFT.

We note that via Myhill-type constructions, general finite state Markov processes can be viewed as evolving on semigroups [12-14], and thus the issue of the complexity of semigroup algebra multiplication is important in understanding the complexity of the filtering problem. Work in this area is continuing [17].

4. Large-Scale Systems with Group-Symmetric Interconnections

The results described in this section are preliminary in nature and a more detailed treatment will be given in the forthcoming report [18]. We also refer the reader to results in [19], [20] that are along these lines. The basic idea of this work is to consider control and estimation for systems that are made up of many interconnected subsystems. The major problems that one is interested in are either

(1) Algorithmic -- i.e. the systems are so complex that the usual design and analysis algorithms (computer solutions of Lyapunov or Riccati equations, etc.) cannot handle them. One wants to devise special algorithms for analysis and synthesis that utilize the system structure.

or (2) Implementation -- i.e. often the usual system designs cannot be implemented in real time, and one wants to utilize system structure to speed up implementation (for example by taking advantage of sparseness) or to determine faster, perhaps decentralized, designs.

As in the preceding sections, our approach consists of examining systems that possess certain structural properties that we can exploit. Consider the dynamical system

$$x(k+1) = Ax(k) + Bu(k) + w(k) \tag{4.1}$$

$$y(k) = Cx(k) + v(k) \tag{4.2}$$

and suppose we want to design a feedback controller that produces $u(k)$ based on $y(j)$, $j \leq k$ and that minimizes

$$J = E\left\{\sum_{k=0}^{\infty}[x'(k)Qx(k) + u'(k)Ru(k)]\right\} \tag{4.3}$$

The solution to this problem is well-known and it consists of a Kalman filter to estimate $x(k)$ followed by a calculation of $u(k)$. The design involves the solution of 2 Riccati equations of dimension $n=\dim(x)$. If n is large, this is an extremely difficult problem. In addition, the on-line implementation also requires a large number of computations per stage.

Suppose, however, that the overall system can be decoupled into smaller subsystems by change of basis on the state, input, and output spaces. Then, we can solve lower-dimensional problems and things become simpler. The difficulties are the following

(1) In general it is not easy to find such basis transformations.

(2) In general one can't expect to decouple everything (including the cost J).

(3) For on-line implementation one would have to perform the input and output transformations in real time. In general this is a difficult computational

task.

Consider, however, the following situation: our system is made up of N identical subsystems numbered from 0 through N-1, and they are interconnected in a circularly-symmetric fashion. That is, x, u, y, w, and v are composed of subvectors for each subsystem, and if we use a subscript to denote the subsystem number, we have

$$x_i(k+1) = \sum_{j=0}^{N-1} A_j x_{i-j}(k) + \sum_{j=0}^{N-1} B_j u_{i-j}(k) + w_i(k) \tag{4.4}$$

$$y_i(k) = \sum_{j=0}^{N-1} C_j x_{i-j}(k) + v_i(k) \tag{4.5}$$

(note that in this case A in (4.1) is block circulant). Suppose also that

$$E[w_i(k)w_j'(\ell)] = \Sigma_{i-j}\delta_{k\ell} \tag{4.6}$$

$$E[v_i(k)v_j'(\ell)] = \Theta_{i-j}\delta_{k\ell} \tag{4.7}$$

and

$$J=E\sum_{k=0}^{\infty}\sum_{i,j=0}^{N-1}[x_i'(k)Q_{i-j}x_j(k)+u_i'(k)R_{i-j}u_j(k)] \tag{4.8}$$

Let us then take the <u>subsystem transforms</u>

$$x(k,\ell) = \sum_{i=0}^{N-1} x_i(k)\gamma^{-i\ell}$$

$$A(\ell) = \sum_{i=0}^{N-1} A_i\gamma^{-i\ell} \tag{4.9}$$

etc.

One then finds the following transformed problem (here * = complex conjugate):

$$x(k+1,\ell) = A(\ell)x(k,\ell) + B(\ell)u(k,\ell) + w(k,\ell) \tag{4.10}$$

$$y(k,\ell) = C(\ell)x(h,\ell) + v(k,\ell) \tag{4.11}$$

where

$$E(w(k,\ell)w^*(j,m)') = \Sigma(\ell)\delta_{kj}\,\delta_{\ell m} \tag{4.12}$$

$$E(v(k,\ell)v^*(j,m)') = \Theta(\ell)\delta_{kj}\,\delta_{\ell m} \tag{4.13}$$

$$J = \frac{1}{N}\left\{\sum_{k=0}^{\infty}\sum_{\ell=0}^{N-1} x^*(k,\ell)'Q(\ell)x(k,\ell)+u^*(k,\ell)'R(\ell)u(k,\ell)\right\} \tag{4.14}$$

Thus, we have reduced the problem to N decoupled problems. These are complex-valued problems, but since x is real

$$A(-\ell) = A^*(\ell), \text{ etc.} \tag{4.15}$$

and we need only solve the problems for

$$\ell=0,1,\ldots,\left\lfloor\frac{N-1}{2}\right\rfloor .$$

For this problem we have reduced the algorithmic complexity (by obtaining lower order problems to be solved) <u>and</u> the complexity of implementation. That is, we implement the system as follows: (1) obtain measurements $y_i(k)$; (2) Use FFT to compute $y(k,\ell)$; (3) process this data through $\frac{N-1}{2}$ decoupled Kalman filters and optimal gains to obtain $u(k,\ell)$; and (4) use the inverse FFT to obtain $u(k)$.

The above analysis barely scratches the surface of what can be done with such systems. Other issues that we have considered and are studying presently are:

(1) Fast algorithms for stability tests, Lyapunov and Riccati equations, pole placement, etc.

(2) Extensions to other types of group symmetries (with the use of the corresponding transforms as described in the preceding section). Also the possible (simpler) use of number-theoretic transforms.

(3) Decentralized control (viewed as a type of "subsystem low-pass filter").

(4) Utility of such control laws when the system isn't quite of this form.

Results in these and related areas will be reported in [18].

References

1. J.T. Lo and A.S. Willsky, "Estimation for Rotational Processes with One Degree of Freedom I: Introduction and Continuous Time Processes," <u>IEEE Trans. Aut. Contr.</u>, Vol. AC-20, No. 1, Feb. 1975, pp. 10-21.

2. A.S. Willsky and J.T. Lo, "Estimation for Rotational Processes with One Degree of Freedom II and III," <u>IEEE Trans. Aut. Contr.</u>, Vol. AC-20, No. 1, Feb. 1975, pp. 22-33.

3. J.T. Lo and A.S. Willsky, "Stochastic Control of Rotational Processes with One Degree of Freedom," <u>SIAM J. Contr.</u>, Vol. 13, No. 4, Nov. 1975, pp. 886-898.

4. A.S. Willsky, S.I. Marcus, and D.N. Martin, "On the Stochastic Stability of Linear Systems Containing Colored Multiplicative Noise," <u>IEEE Trans. Aut. Contr.</u>, Vol. AC-20, Oct. 1975, pp. 711-713.

5. A.S. Willsky and S.I. Marcus, "Analysis of Bilinear Noise Models in Circuits and Devices," J. Franklin Inst., Vol. 301, Nos. 1 and 2, Jan-Feb. 1976, pp.103-122.

6. S.I. Marcus, "Estimation and Analysis of Nonlinear Stochastic Systems," Rept. ESL-R-601, M.I.T. Electronic Systems Lab., Cambridge, Mass., June 1975.

7. A.S. Willsky, "Estimation and Detection of Signals in Multiplicative Noise," IEEE Trans. Inf. Th., Vol. IT-21, July 1975, pp. 472-474.

8. S.I. Marcus and A.S. Willsky, "Algebraic Structure and Finite Dimensional Nonlinear Estimation," Proc. Symp. on Alg. Sys. Th., Udine, Italy, June 1975; also submitted to SIAM J. Math. Analysis.

9. R.W. Brockett, "System Theory on Group Manifolds and Coset Spaces," SIAM J. Control, Vol. 10, No.2, May 1972, pp.265-284.

10. D.Q. Mayne and R.W. Brockett, eds., Geometric Methods in System Theory, Reidel Pub. Co., The Netherlands, 1973.

11. R.W. Brockett, "Volterra Series and Geometric Control Theory," Proc. 1975 IFAC, Cambridge, Mass., Aug. 1975.

12. A.S. Willsky, "A Finite Fourier Transform Approach to Estimation on Cyclic Groups," Proc. Fifth Symp. on Nonlin. Est. and Appl. San Diego, Calif., Sept. 1974.

13. A.S. Willsky, "Dynamical Systems Defined on Groups: Structural Properties and Estimation," Ph.D. thesis, Dept. of Aero. and Astro., M.I.T., Cambridge, Mass., June 1973.

14. A.S. Willsky, "Filtering for Random Finite Group Homomorphic Sequential Systems," Proc. Symp. on Alg. Sys. Th., Udine, Italy, June 1975.

15. A.S. Willsky, "On the Algebraic Structure of Certain Partially Observable Finite State Markov Processes," to appear.

16. C.W. Curtis and I. Reiner, Representation Theory of Finite Groups and Associative Algebras, Interscience, New York, 1966.

17. M. Loui, "Complexity of Multiplication in Semisimple Algebras," S.M. thesis, Dept. of Elec. Eng. and Comp. Sci., M.I.T., Camb., Mass., Jan. 1977.

18. S. Attasi, N.R. Sandell, J. Wall, and A.S. Willsky, "Control and Estimation for Circulant Systems," to appear.

19. S.M. Melzer and B.C. Kuo, "Optimal Regulation of Systems Described by a Countably Infinite Number of Objects," Automatica, Vol. 7, 1971, pp. 359-366.

20. K.C. Chiu, "Optimal Decentralized Regulation for a String of Coupled Systems," IEEE Trans. Aut. Contr., Vol. AC-19, June 1974, pp. 243-246.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 18 AFOSR - TR-76-1309	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) 6 THE USE OF GROUP-THEORETIC CONCEPTS IN SOLVING PROBLEMS IN ESTIMATION AND CONTROL		5. TYPE OF REPORT & PERIOD COVERED Interim
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Alan S. Willsky 10		8. CONTRACT OR GRANT NUMBER(s) 15 AF-AFOSR -2273-72
9. PERFORMING ORGANIZATION NAME AND ADDRESS Massachusetts Institute of Technolcoy Department of Electrical Engineering & Computer Sci Cambridge, MA 02139		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 16 61102F 2304 A1 17
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Office of Scientific Research/NM Bolling AFB, Washington, DC 20332		12. REPORT DATE 11 1976
		13. NUMBER OF PAGES 8
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) 12 9p.		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

In this paper we describe some results obtained over the past few years in the areas of estimation and control. The point of this research is to find sub-classes of systems that possess sufficient structure to allow us to determine optimal solutions or computationally feasible algorithms. In particular, we utilize several results from group theory (both for Lie and finite groups) to obtain our results.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED 409 490

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)